

вычислительным алгоритмам является обеспечение стабильности по отношению к численным шумам при сохранении разумного потребления компьютерных ресурсов.

Одним из способов удовлетворения приведённых выше условий состоит в развитии адекватных упрощённых моделей, таких как 2d3v код LCODE [2]. Основой численного алгоритма данного кода является метод частиц в ячейках, в котором частицам плазмы соответствует набор макрочастиц, а электромагнитные поля, плотности заряда и тока представляются на сетке. Кроме того, в коде LCODE используется квазистатическое приближение. Данное приближение основано на существовании большой относительной разницы характерных времён эволюции для плазмы и драйвера. Следуя квазистатическому приближению, на одном шаге по времени драйвер можно считать неизменным и исследовать плазменную динамику в создаваемых им полях.

На данный момент разрабатывается код LCODE-3D. По сравнению с его предшественником, двумерным кодом LCODE, трёхмерный код позволяет подробнее моделировать эксперименты и точнее исследовать физические явления. Тем не менее, необходимы дополнительные усилия для повышения стабильности по отношению к численным шумам. В кодах, основанных на методе частиц в ячейках, возникающие шумы могут иметь природу, связанную как с численным решением уравнений Максвелла, так и с «саморазогревом» плазмы. Первостепенной задачей в борьбе с шумами является определение параметров, характеризующих устойчивость решения и позволяющих сравнивать эффективность различных методов подавления шумов.

В данной работе рассмотрены методы подавления и фильтрации численных шумов для кода LCODE-3D. Определены параметры эффективности для данных методов. На примере тестовой задачи, имеющей аналитическое решение, найден оптимальный метод стабилизации.

Список публикаций:

- [1] Костюков И. Ю., Пухов А. М. Плазменные методы ускорения электронов: современное состояние и перспективы // УФН, т.85, №1, с.89-96 (2015).  
[2] Lotov K. V. Fine wakefield structure in the blowout regime of plasma wakefield accelerators // Phys. Rev. ST - Accel. Beams 6 (2003), p.061301.

## **О применении метода анализа семантических сетей в медицинской статистике**

**Кушеева Мария Николаевна**

*Аюшеев Донир Цыденович*

*Национальный исследовательский Томский политехнический университет*

*Сибирский государственный медицинский университет*

*Аксёнов Сергей Владимирович, к.т.н.*

[marykush98@mail.ru](mailto:marykush98@mail.ru)

В настоящее время текстовые медицинские данные, такие как анамнез, результаты осмотров и обследований, хранятся в неструктурированном виде. Проблемам создания систем обработки естественного языка, представленных в текстовом виде, посвящены работы огромного количества исследователей (Баранов А.А., Намазова-Баранова Л.С., Смирнов И.В., Девяткин Д.А., Шелманов А.О.) [1,2]. Анализ их работ показал, что семантические сети также могут быть применимы в данной области и имеют огромный потенциал в использовании. Поскольку методам их построения посвящено уже много научных работ, хотелось бы обратить внимание на то, как именно семантические сети могут быть применены в структуризации медицинских данных, улучшении работы медицинского персонала и увеличении качества оказываемых медицинских услуг. Таким образом, данное исследование является достаточно актуальным, потому что метод анализа семантических сетей текстов в перспективе должен упростить задачу анализа медицинских данных, подняв медицинскую статистику и даже диагностику на новый уровень.

Итак, *объектом исследования* являются семантические сети, а *предметом исследования* – метод анализа семантических сетей.

Основной целью исследования можно считать изучение способа применения семантических сетей в обработке и классификации медицинских данных для медицинской статистики.

Суть метода анализа семантических сетей заключается в сравнении семантических сетей, построенных на основе историй болезней пациентов, друг с другом. Итак, имея в распоряжении данные о половозрастной категории пациента, список его жалоб, а также объективный и локальный статус осмотра, необходимо построить семантическую сеть, которая впоследствии будет сравниваться с другими семантическими сетями,

построенным по данным пациентов, диагноз и протекание болезни которых уже известны. Естественно, что такие семантические сети, уже будут разделены на группы по диагнозам.

В настоящее время существует метод оценки качества семантических сетей [3], суть которого заключается в выведении коэффициента совпадения одной семантической сети, называемой эталонной, с другой сетью, называемой кандидатом. Используя этот метод, но принимая в качестве эталонной сети сеть, построенную по истории с известным диагнозом, а в качестве сети-кандидата сеть, построенную по истории, не имеющих таких данных, можно вычислить некоторый коэффициент совпадения.

Далее для коэффициентов сравнения с семантическими сетями, принадлежащими одной группе, рассчитываем среднее арифметическое значение. Т.к. имеется множество различных групп семантических сетей, разбитых по диагнозам, в конечном итоге мы получаем множество средних значений коэффициентов совпадения сети-кандидата с различными группами. Данная совокупность значений итоговых коэффициентов позволит сделать вывод о том, насколько в процентном отношении рассматриваемый пациент может иметь тот или иной диагноз.

Побочным же результатом данного исследования также может являться ранжирование признаков определенной болезни. Т.к. в конечном итоге мы имеем разделенные на группы семантические сети, при объединении в каждой группе семантических сетей в одну и посчитав веса терминов, можно будет судить о важности и встречаемости определенного признака для каждого типа болезни.

Однако, в настоящий момент данное исследование нуждается в качественно проведенных экспериментах, для проведения которых необходимо создать программный прототип, сделать выборку историй болезни, и построить соответствующие им семантические сети. Исследование и разработка метода внесут свой вклад в автоматизацию и компьютеризацию медицинской статистики, но данный метод все равно будет иметь некоторую погрешность, которую также необходимо вычислить в ходе экспериментов.

Список публикаций:

- [1] Баранов А.А., Намазова-Баранова Л.С., Смирнов И.В. и др. Технологии комплексного интеллектуального анализа клинических данных // Вестник Российской академии медицинских наук. – 2016. – с. 160-171
- [2] Лапаев М.В. Система обработки текстовых медицинских данных // Известия СПбГЭТУ «ЛЭТИ». – 2016. – № 9. – с. 11-15
- [3] Аюшеева Н.Н., Гомбожапова Т.Н., Кушеева М.Н. Метод оценки качества семантических сетей // Современные наукоемкие технологии. – 2019. – № 11 (часть 1) – с. 77-81

## **Исследование реализации алгоритмов обработки дифрактограмм на основе кратномасштабного анализа в программной среде Matlab**

**Марудов Андрей Алексеевич**

**Сидорова Екатерина Игоревна**

*Ярославский государственный университет им. П.Г. Демидова*

*Московский Сергей Борисович*

[marudov.andrey@gmail.com](mailto:marudov.andrey@gmail.com)

Характерной особенностью задач, требующих обработки значительного объема данных, являются большая множественность событий и высокий уровень шума. Внедрение методов вейвлет-анализа в алгоритмы обработки данных наглядно показывает их способность комплексно подходить к решению задач. Одним из методов обработки зашумленного сигнала является трешолдинг (ТШ). Он заключается в вейвлет-декомпозиции сигнала до уровня  $j$  и удалении достаточно маленьких вейвлет-коэффициентов на каждом уровне разложения.

В качестве материала для обработки были взяты данные образцов пленок V-O, полученные на дифрактометре ARL X'TRA (Thermo Scientific - Швейцария, 2009), находящемся в ЦКП ДМНС.

Задачей является численное исследование влияния параметров трешолдинга и способов пороговой обработки на качество шумоподавления. Основным критерий оценки – относительное среднеквадратичное отклонение шума. Всего было рассмотрено три различных алгоритма, представленных в пакете расширения Wavelet Toolbox компьютерной системы MATLAB.

Первый и самый простой из них – алгоритм ТШ с использованием функции, которая находит значения параметров по умолчанию для всех общих процедур, связанных с удалением шума, используя вейвлеты. Параметры по умолчанию следующие:

- глобальный порог  $thr = \sqrt{2 \ln L} \cdot E$ , где  $L$  – длина сигнала,  $E$  – уровень шума;
- сохранение аппроксимирующих коэффициентов;